PLOS ONE

RESEARCH ARTICLE

A fast combination method in DSmT and its application to recommender system

Yilin Dong[1☯], **Xinde Li**[1☯*], **Yihai Liu**[2]

1 Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing, Jiangsu Province, China, **2** Jiangsu Automation Research Institute, Lianyungang, Jiangsu Province, China

☯ These authors contributed equally to this work.
* xindeli@seu.edu.cn

Citation: Dong Y, Li X, Liu Y (2018) A fast combination method in DSmT and its application to recommender system. PLoS ONE 13(1): e0189703. https://doi.org/10.1371/journal.pone.0189703

Editor: Yong Deng, Southwest University, CHINA

Received: August 28, 2017

Accepted: November 30, 2017

Published: January 19, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files. The two datasets, Movielens and Flixster, are available on Figshare at DOIs 10.6084/m9.figshare.5677750 and 10.6084/m9.figshare.5677741 respectively.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61573097, 91748106, in part by Key Laboratory of Integrated Automation of Process Industry (PAL-N201704), in part by the Qing Lan Project and Six Major Top-talent Plan, and in part by the Priority Academic Program Development of

Abstract

In many applications involving epistemic uncertainties usually modeled by belief functions, it is often necessary to approximate general (non-Bayesian) basic belief assignments (BBAs) to subjective probabilities (called Bayesian BBAs). This necessity occurs if one needs to embed the fusion result in a system based on the probabilistic framework and Bayesian inference (e.g. tracking systems), or if one needs to make a decision in the decision making problems. In this paper, we present a new fast combination method, called modified rigid coarsening (MRC), to obtain the final Bayesian BBAs based on hierarchical decomposition (coarsening) of the frame of discernment. Regarding this method, focal elements with probabilities are coarsened efficiently to reduce computational complexity in the process of combination by using disagreement vector and a simple dichotomous approach. In order to prove the practicality of our approach, this new approach is applied to combine users' soft preferences in recommender systems (RSs). Additionally, in order to make a comprehensive performance comparison, the proportional conflict redistribution rule #6 (PCR6) is regarded as a baseline in a range of experiments. According to the results of experiments, MRC is more effective in accuracy of recommendations compared to original Rigid Coarsening (RC) method and comparable in computational time.

Introduction

The theory of belief functions, known as Dempster-Shafer Theory (DST) was developed by Shafer [1] in 1976 from Dempster's works [2]. Belief functions allow one to model epistemic uncertainty [3] and they have been already used in many applications since the 1990's [4], mainly those relevant to expert systems, decision-making support and information fusion. To palliate some limitations (such as high computational compelxity) of DST, Dezert and Smarandache proposed an extended mathematical framework of belief functions with new efficient quantitative and qualitative rules of combinations, which was called DSmT (Dezert and Smarandache Theory) in literature [5, 6] with applications listed in [7]. One of the major drawbacks of DST and DSmT is their high computational complexities, on condition that the fusion

Jiangsu Higher Education Institutions to XL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

space (i.e. frame of discernment—FoD) and the number of sources to combine are large. DSmT is more complex than DST, and the Proportional Conflict Redistribution rule #6 (PCR6 rule) becomes computationally intractable in the worst case as soon as the cardinality of the Frame of Discernment (FoD) is greater than six.

To reduce the computational cost of operations with belief functions when the number of focal elements is very large, several approaches have been proposed by different authors. Basically, the existing approaches rely either on efficient implementations of computations as proposed for instance in [8, 9], or on approximation techniques of original Basic Belief Assignment (BBA) to combine [10–14], or both. From a fusion standpoint, two approaches are usually adopted: 1) one can approximate at first each BBA in subjective probabilities and use Bayes fusion rule to get the final Bayesian BBA [11, 12], or 2) one can fuse all the BBAs with a fusion rule, typically Dempster-Shafer's, or proportional conflict redistribution rule #6 (PCR6) rules (which is very costly in computations), and convert the combined BBA in a subjective probability measure [10, 14]. The former method is the simplest method but it generates a high loss of information included in the original BBAs, whereas the latter method is intractable for high dimension issues.

This paper presents a new combination method, called modified rigid coarsening (MRC), to get the final Bayesian BBAs based on hierarchical decomposition (coarsening) of the frame of discernment, which can be seen as an intermediary approach between the two aforementioned methods. This hierarchical structure allows to encompass bintree decomposition and mass of coarsening FoD on it. To prove the practicality of our proposed method, MRC is applied to combine users' preferences so as to provide the suitable recommendation for RSs. Preliminary work on original rigid coarsening (RC) has been published in our recent work [15] (This is an extended version of the paper presented at the 20th IEEE International Conference on Information Fusion, XIAN, China). In this paper, more detailed analyses of this new combination method are provided. More importantly, this innovative method is also applied into the real application. These are all added values (contributions) of this paper.

The main contributions of this paper are:

1. the presentation of the FoD bintree decomposition on which will be done the BBAs approximations;
2. user preferences in Recommender Systems (RSs) are modeled by DSmT-Modeling Function.

In order to measure the efficiency and effectiveness of the MRC, it is integrated in the RSs based on DSmT and compared to traditional methods in the experiments. The results show that regarding the accuracy of recommendations, MRC is extremely close to classical PCR6; and the computational time of MRC can be obviously superior to that of PCR6.

The remainder of this paper is organized as follows. In section 2, we review relevant prior work on DST and DSmT first. In section 3, MRC is presented. In section 4, a recommendation system based on DSmT, that employs MRC to combine users' preferences, is shown. In section 5, we evaluate our proposed algorithm based on two public datasets: Movielens and Flixster. Finally, we conclude and discuss future work.

Mathematical background

This section provides a brief reminder of the basics of DST and DSmT, which is necessary for the presentation and understanding of the more general MRC of Section 3.

In DST framework, the frame of discernment (Here, we use the symbol $\triangleq$ to mean equals by definition.) $\Theta \triangleq \{\theta_1, \ldots, \theta_n\}$ $(n \geq 2)$ is a set of exhaustive and exclusive elements

(hypotheses) which represents the possible solutions of the problem under consideration and thus Shafer's model assumes $\theta_i \cap \theta_j = \emptyset$ for $i \neq j$ in $\{1, \ldots, n\}$. A basic belief assignment (BBA) $m(\cdot)$ is defined by the mapping: $2^\Theta \mapsto [0, 1]$, verifying $m(\emptyset) = 0$ and $\sum_{A \in 2^\Theta} m(A) = 1$. In DSmT, one can abandon Shafer's model (if Shafer's model doesn't fit with the problem) and refute the principle of the third excluded middle. The third excluded middle principle assumes the existence of the complement for any elements/propositions belonging to the power set 2^Θ. Instead of defining the BBAs on the power set $2^\Theta \triangleq (\Theta, \cup)$ of the FoD, the BBAs are defined on the so-called *hyper-power set* (or Dedekind's lattice) denoted $D^\Theta \triangleq (\Theta, \cup, \cap)$ whose cardinalities follows Dedekind's numbers sequence, see [6], Vol.1 for details and examples. A (generalized) BBA, called a mass function, $m(\cdot)$ is defined by the mapping: $D^\Theta \mapsto [0, 1]$, verifying $m(\emptyset) = 0$ and $\sum_{A \in D^\Theta} m(A) = 1$. The DSmT framework encompasses DST framework because $2^\Theta \subset D^\Theta$. In DSmT, we can take into account also a set of *integrity constraints* on the FoD (if known), by specifying all the pairs of elements which are really disjoint. Stated otherwise, Shafer's model is a specific DSm model where all elements are deemed to be disjoint. $A \in D^\Theta$ is called a focal element of $m(.)$ if $m(A) > 0$. A BBA is called a Bayesian BBA if all of its focal elements are singletons and Shafer's model is assumed, otherwise it is called non-Bayesian [1]. A full ignorance source is represented by the vacuous BBA $m_v(\Theta) = 1$. The belief (or credibility) and plausibility functions are respectively defined by $Bel(X) \triangleq \sum_{Y \in D^\Theta | Y \subseteq X} m(Y)$ and $Pl(X) \triangleq \sum_{Y \in D^\Theta | Y \cap X \neq \emptyset} m(Y)$. $BI(X) \triangleq [Bel(X), Pl(X)]$ is called the belief interval of X. Its length $U(X) \triangleq Pl(X) - Bel(X)$ measures the degree of uncertainty of X.

In 1976, Shafer did propose Dempster's rule and we use DS index to refer to Dempster-Shafer's rule (DS rule) because Shafer did really promote Dempster's rule in in his milestone book [1]) to combine BBAs in DST framework. DS rule is defined by $m_{DS}(\emptyset) = 0$ and $\forall A \in 2^\Theta \setminus \{\emptyset\}$,

$$m_{DS}(A) = \frac{\sum_{B,C \in 2^\Theta | B \cap C = A} m_1(B) m_2(C)}{1 - \sum_{B,C \in 2^\Theta | B \cap C = \emptyset} m_1(B) m_2(C)} \tag{1}$$

The DS rule formula is commutative and associative and can be easily extended to the fusion of $S > 2$ BBAs. Unfortunately, DS rule has been highly disputed during the last decades by many authors because of its counter-intuitive behavior in high or even low conflict situations, and that is why many rules of combination were proposed in literature to combine BBAs [16]. To palliate DS rule drawbacks, the very interesting PCR6 was proposed in DSmT and it is usually adopted (PCR6 rule coincides with PCR5 when combining only two BBAs [6]) in recent applications of DSmT. The fusion of two BBAs $m_1(.)$ and $m_2(.)$ by the PCR6 rule is obtained by $m_{PCR6}(\emptyset) = 0$ and $\forall A \in D^\Theta \setminus \{\emptyset\}$

$$\begin{aligned} m_{PCR6}(A) &= m_{12}(A) \\ &\quad + \sum_{B \in D^\Theta \setminus \{A\} | A \cap B = \emptyset} \left[\frac{m_1(A)^2 m_2(B)}{m_1(A) + m_2(B)} + \frac{m_2(A)^2 m_1(B)}{m_2(A) + m_1(B)} \right] \end{aligned} \tag{2}$$

where $m_{12}(A) = \sum_{B,C \in D^\Theta | B \cap C = A} m_1(B) m_2(C)$ is the conjunctive operator, and each element A and B are expressed in their disjunctive normal form. If the denominator involved in the fraction is zero, then this fraction is discarded. The general PCR6 formula for combining more than two BBAs altogether is given in [6], Vol. 3. We adopt the generic notation $m_{12}^{PCR6}(.) = PCR6(m_1(.), m_2(.))$ to denote the fusion of $m_1(.)$ and $m_2(.)$ by PCR6 rule. PCR6 is not associative and PCR6 rule can also be applied in DST framework (with Shafer's model of FoD) by replacing D^Θ by 2^Θ in Eq (2).

Modified rigid coarsening for fusion of Bayesian BBAs

Here, we introduce the principle of MRC of FoD to reduce the computational complexity of PCR6 combination of original Bayesian BBAs. Considering the case of non-Bayesian BBAs, it requires decoupling all non-singletons in these BBAs in advance, The fusion of original non-Bayesian BBAs needs to be decoupled by using DSmP in advance, which will explain in Section 4.

Rigid coarsening

This proposal was initially called *rigid coarsening (RC)* in our previous works [17–19] and currently improved in our recent work [15]. The goal of this coarsening is to replace the original (refined) FoD Θ by a set of coarsened ones to make computation of the PCR6 rule tractable. Because we consider here only Bayesian BBA to combine, their focal elements are only singletons of the FoD $\Theta \triangleq \{\theta_1, \ldots, \theta_n\}$, with $n \geq 2$, and we assume Shafer's model of the FoD Θ. A coarsening of the FoD Θ means to replace it with another FoD less specific of smaller dimension $\Omega = \{\omega_1, \ldots, \omega_k\}$ with $k < n$ from the elements of Θ. This can be done in many ways depending the problem under consideration. Generally, the elements of Ω are singletons of Θ, and disjunctions of elements of Θ. For example, if $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, then a possible coarsened frame built from Θ could be, for instance, $\Omega = \{\omega_1 = \theta_1, \omega_2 = \theta_2, \omega_3 = \theta_3 \cup \theta_4\}$, or $\Omega = \{w_1 = \theta_1 \cup \theta_2, \omega_2 = \theta_3 \cup \theta_4\}$, etc.

Definition 1: *When dealing with Bayesian BBAs, the projection (For clarity and convenience, we put explicitly as upper index the FoD for which the belief mass refers) $m^{\Omega}(.)$ of the original BBA $m^{\Theta}(.)$ is simply obtained by taking*

$$m^{\Omega}(\omega_i) = \sum_{\theta_j \subseteq \omega_i} m^{\Theta}(\theta_j) \tag{3}$$

The rigid coarsening process is a simple dichotomous approach of coarsening obtained as follows:

- If $n = |\Theta|$ is an even number:
  The disjunction of the $n/2$ first elements θ_1 to $\theta_{\frac{n}{2}}$ of Θ define the element ω_1 of Ω, and the last $n/2$ elements $\theta_{\frac{n}{2}+1}$ to θ_n of Θ define the element ω_2 of Ω, that is

  $$\Omega \triangleq \{\omega_1 = \theta_1 \cup \ldots \cup \theta_{\frac{n}{2}}, \omega_2 = \theta_{\frac{n}{2}+1} \cup \ldots \cup \theta_n\}$$

  and based on Eq (3), one has

  $$m^{\Omega}(\omega_1) = \sum_{j=1,\ldots,\frac{n}{2}} m^{\Theta}(\theta_j) \tag{4}$$

  $$m^{\Omega}(\omega_2) = \sum_{j=\frac{n}{2}+1,\ldots,n} m^{\Theta}(\theta_j) \tag{5}$$

  For example, if $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, and one considers the Bayesian BBA $m^{\Theta}(\theta_1) = 0.1$, $m^{\Theta}(\theta_2) = 0.2$, $m^{\Theta}(\theta_3) = 0.3$ and $m^{\Theta}(\theta_4) = 0.4$, then $\Omega = \{\omega_1 = \theta_1 \cup \theta_2, \omega_2 = \theta_3 \cup \theta_4\}$ and $m^{\Omega}(\omega_1) = 0.1 + 0.2 = 0.3$ and $m^{\Omega}(\omega_2) = 0.3 + 0.4 = 0.7$.

- If $n = |\Theta|$ is an odd number:
  In this case, the element ω_1 of the coarsened frame Ω is the disjunction of the $[n/2 + 1]$ (The notation $[x]$ means the integer part of x) first elements of Θ, and the element ω_2 is the

Table 1. Three Bayesian BBAs for Example 1.

Focal elem.	$m_1^{\Theta}(.)$	$m_2^{\Theta}(.)$	$m_3^{\Theta}(.)$
θ_1	0.1	0.4	0
θ_2	0.2	0	0.1
θ_3	0.3	0.1	0.5
θ_4	0.3	0.1	0.4
θ_5	0.1	0.4	0

https://doi.org/10.1371/journal.pone.0189703.t001

disjunction of other elements of Θ. That is

$$\Omega \triangleq \{\omega_1 = \theta_1 \cup \ldots \cup \theta_{\left[\frac{n}{2}+1\right]}, \omega_2 = \theta_{\left[\frac{n}{2}+1\right]+1} \cup \ldots \cup \theta_n\}$$

and based on Eq (3), one has

$$m^{\Omega}(\omega_1) = \sum_{j=1,\ldots,\left[\frac{n}{2}+1\right]} m^{\Theta}(\theta_j) \tag{6}$$

$$m^{\Omega}(\omega_2) = \sum_{j=\left[\frac{n}{2}+1\right]+1,\ldots,n} m^{\Theta}(\theta_j) \tag{7}$$

For example, if $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$, and one considers the Bayesian BBA $m^{\Theta}(\theta_1) = 0.1$, $m^{\Theta}(\theta_2) = 0.2$, $m^{\Theta}(\theta_3) = 0.3$, $m^{\Theta}(\theta_4) = 0.3$ and $m^{\Theta}(\theta_5) = 0.1$, then $\Omega = \{\omega_1 = \theta_1 \cup \theta_2 \cup \theta_3, \omega_2 = \theta_4 \cup \theta_5\}$ and $m^{\Omega}(\omega_1) = 0.1 + 0.2 + 0.3 = 0.6$ and $m^{\Omega}(\omega_2) = 0.3 + 0.1 = 0.4$.

Of course, the same coarsening strategy applies to all original BBAs $m_s^{\Theta}(.)$, $s = 1, \ldots S$ of the $S > 1$ sources of evidence to work with less specific BBAs $m_s^{\Omega}(.)$, $s = 1, \ldots S$. The less specific BBAs (called coarsened BBAs by abuse of language) can then be combined with the PCR6 rule of combination according to formula Eq (2). This dichotomous coarsening method is repeated iteratively l times as schematically represented by a bintree. Here, we consider bintree only for simplicity, which means that the coarsened frame Ω consists of two elements only. Of course a similar method can be used with tri-tree, quad-tree, etc. The last step of this hierarchical process is to calculate the combined (Bayesian) BBA of all focal elements according to the connection weights of the bintree structure, where the number of layers l of the tree depends on the cardinality |Θ| of the original FoD Θ. Specifically, the mass of each focal element is updated depending on the connection weights of link paths from root to terminal nodes. This principle is illustrated in details in the following example.

Example 1: Let's consider $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$, and the following three Bayesian BBAs can be seen in Table 1:

The rigid coarsening and fusion of BBAs is deduced from the following steps:

Step 1: We define the bintree structure based on iterative half split of FoD as shown in Fig 1.

The connecting weights are denoted as $\lambda_1, \ldots, \lambda_8$. The elements of the frames Ω_l are defined as follows:

- At layer $l = 1$: $\Omega_1 = \{\omega_1 \triangleq \theta_1 \cup \theta_2 \cup \theta_3, \omega_2 \triangleq \theta_4 \cup \theta_5\}$
- At layer $l = 2$:

$$\Omega_2 = \{\omega_{11} \triangleq \theta_1 \cup \theta_2, \omega_{12} \triangleq \theta_3, \omega_{21} \triangleq \theta_4, \omega_{22} = \theta_5\}$$

- At layer $l = 3$: $\Omega_3 = \{\omega_{111} \triangleq \theta_1, \omega_{112} \triangleq \theta_2\}$

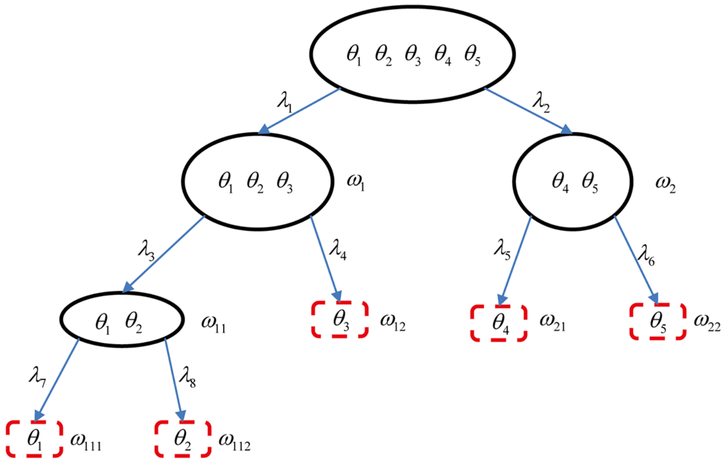


Fig 1. Fusion of Bayesian BBAs using bintree coarsening for Example 1.

https://doi.org/10.1371/journal.pone.0189703.g001

Step 2: The BBAs of elements of the (sub-) frames Ω_l are obtained as follows:

- At layer $l = 1$, we use Eqs (6) and (7) because $|\Theta| = 5$ is an odd number. Therefore, we get the BBAs in Table 2:
- At layer $l = 2$: We work with the two subframes $\Omega_{21} \triangleq \{\omega_{11}, \omega_{12}\}$ and $\Omega_{22} \triangleq \{\omega_{21}, \omega_{22}\}$ of Ω_2 with the BBAs in Tables 3 and 4:
These mass values are obtained by the proportional redistribution of the mass of each focal element with respect to the mass of its parent focal element in the bin tree. For example, $m_2^{\Omega_{21}}(\omega_{11}) = 4/5$ is derived by taking

$$m_2^{\Omega_{21}}(\omega_{11}) = \frac{m_2^{\Theta}(\theta_1) + m_2^{\Theta}(\theta_2)}{m_2^{\Theta}(\theta_1) + m_2^{\Theta}(\theta_2) + m_2^{\Theta}(\theta_3)} = \frac{0.4}{0.5} = \frac{4}{5}$$

Other masses of coarsening focal elements are computed similarly using this proportional redistribution method.

- At layer $l = 3$: We use again the proportional redistribution method which gives us the BBAs of the sub-frames Ω_3 in Table 5:

Step 3: The connection weights λ_i are computed from the assignments of coarsening elements. In each layer l, we fuse sequentially the three BBAs using PCR6 formula Eq (2). Because PCR6 fusion is not associative, we should apply the general PCR6 formula to get best results. Here we use sequential fusion to reduce the computational complexity even if the fusion result is approximate. More precisely, we compute at first $m_{12}^{PCR6,\Omega_l}(.) = PCR6(m_1^{\Omega_l}(.), m_2^{\Omega_l}(.))$ and then $m_{(12)3}^{PCR6,\Omega_l}(.) = PCR6(m_{12}^{PCR6,\Omega_l}(.), m_3^{\Omega_l}(.))$. Hence, we obtain the following connecting weights in the bintree:

Table 2. The BBAs of elements of the sub-frames Ω_1 for Example 1.

Focal elem.	$m_1^{\Theta}(.)$	$m_2^{\Theta}(.)$	$m_3^{\Theta}(.)$
$\omega_1 \triangleq \theta_1 \cup \theta_2 \cup \theta_3$	0.6	0.5	0.6
$\omega_2 \triangleq \theta_4 \cup \theta_5$	0.4	0.5	0.4

https://doi.org/10.1371/journal.pone.0189703.t002

Table 3. The BBAs of elements of the sub-frames Ω_{21} for Example 1.

Focal elem.	$m_1^{\Omega_{21}}(.)$	$m_2^{\Omega_{21}}(.)$	$m_3^{\Omega_{21}}(.)$
$\omega_{11} \triangleq \theta_1 \cup \theta_2$	$\frac{1}{2}$	$\frac{4}{5}$	$\frac{1}{6}$
$\omega_{12} \triangleq \theta_3$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{5}{6}$

https://doi.org/10.1371/journal.pone.0189703.t003

Table 4. The BBAs of elements of the sub-frames Ω_{22} for Example 1.

Focal elem.	$m_1^{\Omega_{22}}(.)$	$m_2^{\Omega_{22}}(.)$	$m_3^{\Omega_{22}}(.)$
$\omega_{21} \triangleq \theta_4$	$\frac{3}{4}$	$\frac{1}{5}$	1
$\omega_{22} \triangleq \theta_5$	$\frac{1}{4}$	$\frac{4}{5}$	0

https://doi.org/10.1371/journal.pone.0189703.t004

Table 5. The BBAs of elements of the sub-frames Ω_3 for Example 1.

Focal elem.	$m_1^{\Omega_3}(.)$	$m_2^{\Omega_3}(.)$	$m_3^{\Omega_3}(.)$
$\omega_{111} \triangleq \theta_1$	$\frac{1}{3}$	1	0
$\omega_{112} \triangleq \theta_2$	$\frac{2}{3}$	0	1

https://doi.org/10.1371/journal.pone.0189703.t005

- At layer $l = 1$:

$$\begin{aligned} \lambda_1 &= m_{(12)3}^{PCR6,\Omega_1}(\omega_1) = 0.6297 \\ \lambda_2 &= m_{(12)3}^{PCR6,\Omega_1}(\omega_2) = 0.3703 \end{aligned}$$

- At layer $l = 2$:

$$\begin{aligned} \lambda_3 &= m_{(12)3}^{PCR6,\Omega_{21}}(\omega_{11}) = 0.4137 \\ \lambda_4 &= m_{(12)3}^{PCR6,\Omega_{21}}(\omega_{12}) = 0.5863 \\ \lambda_5 &= m_{(12)3}^{PCR6,\Omega_{22}}(\omega_{21}) = 0.8121 \\ \lambda_6 &= m_{(12)3}^{PCR6,\Omega_{22}}(\omega_{22}) = 0.1879 \end{aligned}$$

- At layer $l = 3$:

$$\begin{aligned} \lambda_7 &= m_{(12)3}^{PCR6,\Omega_3}(\omega_{111}) = 0.3103 \\ \lambda_8 &= m_{(12)3}^{PCR6,\Omega_3}(\omega_{112}) = 0.6897 \end{aligned}$$

Step 4: The final assignments of elements in original FoD Θ are calculated using the product of the connection weights of link paths from root (top) node to terminal nodes (leaves). We eventually get the *combined and normalized* Bayesian BBA:

$$\begin{aligned} m^{\Theta}(\theta_1) &= \lambda_1 \cdot \lambda_3 \cdot \lambda_7 = 0.6297 \cdot 0.4137 \cdot 0.3103 = 0.0808 \\ m^{\Theta}(\theta_2) &= \lambda_1 \cdot \lambda_3 \cdot \lambda_8 = 0.6297 \cdot 0.4137 \cdot 0.6897 = 0.1797 \\ m^{\Theta}(\theta_3) &= \lambda_1 \cdot \lambda_4 = 0.6297 \cdot 0.5863 = 0.3692 \\ m^{\Theta}(\theta_4) &= \lambda_2 \cdot \lambda_5 = 0.3703 \cdot 0.8121 = 0.3007 \\ m^{\Theta}(\theta_5) &= \lambda_2 \cdot \lambda_6 = 0.3703 \cdot 0.1879 = 0.0696 \end{aligned}$$

Modified rigid coarsening

One of the issues with RC described in the previous section is that *no extra self-information of focal elements is embedded into the coarsening process*. In this paper, the elements θ_i selected to belong to the same group are determined using the consensus information drawn from the BBAs provided by the sources. Specifically, the degrees of disagreement between the provided sources on decisions $(\theta_1, \theta_2, \cdots, \theta_n)$ are first calculated using the belief-interval based distance d_{BI} [20] to obtain *disagreement vector*. And then all focal elements in FoD are sorted in an ascending order. Finally, the simple dichotomous approach is utilized to hierarchical coarsen those **Re-sorted** focal elements.

Calculating the disagreement vector. Let us consider several BBAs $m_s^\Theta(\cdot)$, $(s = 1, \ldots, S)$ defined on same FoD Θ of cardinality $|\Theta| = n$. The specific BBAs $m_{\theta_i}(.)$, $i = 1, \ldots, n$ entirely focused on θ_i are defined by $m_{\theta_i}(\theta_i) = 1$, and for $X \neq \theta_i$ $m_{\theta_i}(X) = 0$.

Definition 2: *The disagreement of opinions of two sources about θ_i is defined as the L_1-distance between the d_{BI} distances of the BBAs $m_s^\Theta(.)$, $s = 1, 2$ to $m_{\theta_i}(.)$, which is expressed by*

$$D_{12}(\theta_i) \triangleq |d_{BI}(m_1^\Theta(\cdot), m_{\theta_i}(\cdot)) - d_{BI}(m_2^\Theta(\cdot), m_{\theta_i}(\cdot))| \tag{8}$$

Definition 3: *The disagreement of opinions of $S \geq 3$ sources about θ_i, is defined as*

$$D_{1-S}(\theta_i) \triangleq \frac{1}{2}\sum_{i=1}^{S}\sum_{j=1}^{S}|d_{BI}(m_i^\Theta(\cdot), m_{\theta_i}(.)) - d_{BI}(m_j^\Theta(\cdot), m_{\theta_i}(.))| \tag{9}$$

where d_{BI} distance is defined by [20] *and proof of Definition 3 is given in* S1 Appendix. *For simplicity, we assume Shafer's model so that $|2^\Theta| = 2^n$, otherwise the number of elements in the summation of* Eq (10) *should be $|D^\Theta| - 1$ with another normalization constant n_c.*

$$d_{BI}^E(m_1, m_2) \triangleq \sqrt{n_c \cdot \sum_{i=1}^{2^n-1}[d^I(BI_1(\theta_i), BI_2(\theta_i))]^2} \tag{10}$$

Here, $n_c = 1/2^{n-1}$ is the normalization constant and $d^I([a, b], [c, d])$ is the Wasserstein's distance defined by $d^I([a,b],[c,d]) = \sqrt{\left[\frac{a+b}{2} - \frac{c+d}{2}\right]^2 + \frac{1}{3}\left[\frac{b-a}{2} - \frac{d-c}{2}\right]^2}$. And $BI(\theta_i) = [Bel(\theta_i), Pl(\theta_i)]$.

The disagreement vector $\mathbf{D}_{1-S}$ is defined by

$$\mathbf{D}_{1-S} \triangleq [D_{1-S}(\theta_1), \ldots, D_{1-S}(\theta_n)] \tag{11}$$

Modified rigid coarsening by using the disagreement vector. Once D_{1-S} is derived, all focal elements $\{\theta_1, \theta_2, \cdots, \theta_n\}$ are sorted according to their corresponding values in D_{1-S}.

Let us revisit example 1 presented in the previous section. It can be verified in applying formula Eq (9) that the disagreement vector $\mathbf{D}_{1-3}$ for this example is equal to

$$\mathbf{D}_{1-3} = [0.4085, 0.2156, 0.3753, 0.2507, 0.4086]$$

The derivation of $D_{1-3}(\theta_1)$ is given below for convenience.

$$\begin{aligned} D_{1-3}(\theta_1) &= |d_{BI}(m_1^\Theta(\cdot), m_{\theta_1}(\theta_1)) - d_{BI}(m_2^\Theta(\cdot), m_{\theta_1}(\theta_1))| \\ &\quad +|d_{BI}(m_2^\Theta(\cdot), m_{\theta_1}(\theta_1)) - d_{BI}(m_3^\Theta(\cdot), m_{\theta_1}(\theta_1))| \\ &\quad +|d_{BI}(m_1^\Theta(\cdot), m_{\theta_1}(\theta_1)) - d_{BI}(m_3^\Theta(\cdot), m_{\theta_1}(\theta_1))| \\ &= 0.4085. \end{aligned}$$

Based on the disagreement vector, a new bintree structure is obtained and shown in Fig 2. Compared with Fig 1, the elements in FoD Θ are grouped more reasonably. In vector $\mathbf{D}_{1-3}$, θ_1

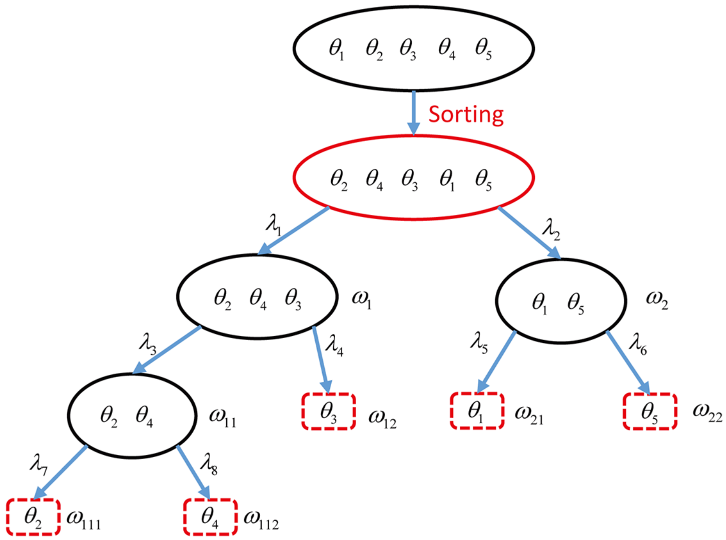


Fig 2. Fusion of Bayesian BBAs using MRC for Example 1.

https://doi.org/10.1371/journal.pone.0189703.g002

and θ_5 lie in similar degree of disagreement so that they are put in the same group. Similarly for θ_2 and θ_4. However, element θ_3 seems *weird*, which is put alone in the process of coarsening. Once this new bintree decomposition is obtained, other steps can be implemented which are identical to rigid coarsening in section to get the final combined BBA.

Step 1: According to Fig 2, the elements of the frames Ω_l are defined as follows:

- At layer $l = 1$: $\Omega_1 = \{\omega_1 \triangleq \theta_2 \cup \theta_4 \cup \theta_3, \omega_2 \triangleq \theta_1 \cup \theta_5\}$
- At layer $l = 2$: $\Omega_2 = \{\omega_{11} \triangleq \theta_2 \cup \theta_4, \omega_{12} \triangleq \theta_3, \omega_{21} \triangleq \theta_1, \omega_{22} \triangleq \theta_5\}$
- At layer $l = 3$: $\Omega_3 = \{\omega_{111} \triangleq \theta_2, \omega_{112} \triangleq \theta_4\}$

Step 2: The BBAs of elements of the (sub-) frames Ω_l are obtained as follows:

- At layer $l = 1$, we use Eqs (6) and (7) and we get (Table 6)
- At layer $l = 2$: We use again the proportional redistribution method which gives us Tables 7 and 8. Here, masses of ω_{21}, ω_{22} in $m_3^{\Omega_{22}}(.)$ are not considered because the mass of their parent focal element ($m_3^{\Omega_1}(\omega_2)$) in bintree is 0.
- At layer $l = 3$: We work with the two subframes $\Omega_3 \triangleq \{\omega_{111}, \omega_{112}\}$ of Ω_3 with the BBAs in Table 9:

Table 6. The BBAs of elements of the sub-frames Ω_1 Using MRC for Example 1.

Focal elem.	$m_1^{\Omega_1}(.)$	$m_2^{\Omega_1}(.)$	$m_3^{\Omega_1}(.)$
$\omega_1 \triangleq \theta_2 \cup \theta_4 \cup \theta_3$	0.8	0.2	1.0
$\omega_2 \triangleq \theta_1 \cup \theta_5$	0.2	0.8	0.0

https://doi.org/10.1371/journal.pone.0189703.t006

Table 7. The BBAs of elements of the sub-frames Ω_{21} Using MRC for Example 1.

Focal elem.	$m_1^{\Omega_{21}}(.)$	$m_2^{\Omega_{21}}(.)$	$m_3^{\Omega_{21}}(.)$
$\omega_{11} \triangleq \theta_2 \cup \theta_4$	$\frac{5}{8}$	$\frac{1}{2}$	$\frac{1}{2}$
$\omega_{12} \triangleq \theta_3$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{2}$

https://doi.org/10.1371/journal.pone.0189703.t007

Table 8. The BBAs of elements of the sub-frames Ω_{22} Using MRC for Example 1.

Focal elem.	$m_1^{\Omega_{22}}(.)$	$m_2^{\Omega_{22}}(.)$	$m_3^{\Omega_{22}}(.)$
$\omega_{21} \triangleq \theta_1$	$\frac{1}{2}$	$\frac{1}{2}$	–
$\omega_{22} \triangleq \theta_5$	$\frac{1}{2}$	$\frac{1}{2}$	–

https://doi.org/10.1371/journal.pone.0189703.t008

Table 9. The BBAs of elements of the sub-frames Ω_3 Using MRC for Example 1.

Focal elem.	$m_1^{\Omega_3}(.)$	$m_2^{\Omega_3}(.)$	$m_3^{\Omega_3}(.)$
$\omega_{111} \triangleq \theta_2$	$\frac{2}{5}$	0.0	$\frac{1}{5}$
$\omega_{112} \triangleq \theta_4$	$\frac{3}{5}$	1.0	$\frac{4}{5}$

https://doi.org/10.1371/journal.pone.0189703.t009

Step 3: The connection weights λ_i are computed from the assignments of coarsening elements. Hence, we obtain the following connecting weights in the bintree:

- At layer $l = 1$:

$$\lambda_1 = 0.8333; \quad \lambda_2 = 0.1667.$$

- At layer $l = 2$:

$$\lambda_3 = 0.5697; \quad \lambda_4 = 0.4303;$$
$$\lambda_5 = 0.5000; \quad \lambda_6 = 0.5000.$$

- At layer $l = 3$:

$$\lambda_7 = 0.0669; \quad \lambda_8 = 0.9331;$$

Step 4: We finally get the following *combined and normalized* Bayesian BBA

$$m^{\Theta}(\cdot) = \{0.0833, 0.0318, 0.3586, 0.4430, 0.0834\}.$$

Summary of the proposed method

The fusion method of BBAs to get a combined Bayesian BBA based on hierarchical decomposition of the FoD consists of several steps (**Algorithm 1**) below illustrated in Fig 3. It is worth noting that when the given BBAs are not Bayesian, the first step is to use the existing Probabilistic Transformation (PT) to transform them to Bayesian BBAs. In order to use the proposed combination method in the RSs, modified rigid coarsening is mathematically denoted as $\bigoplus$ in the following sections.

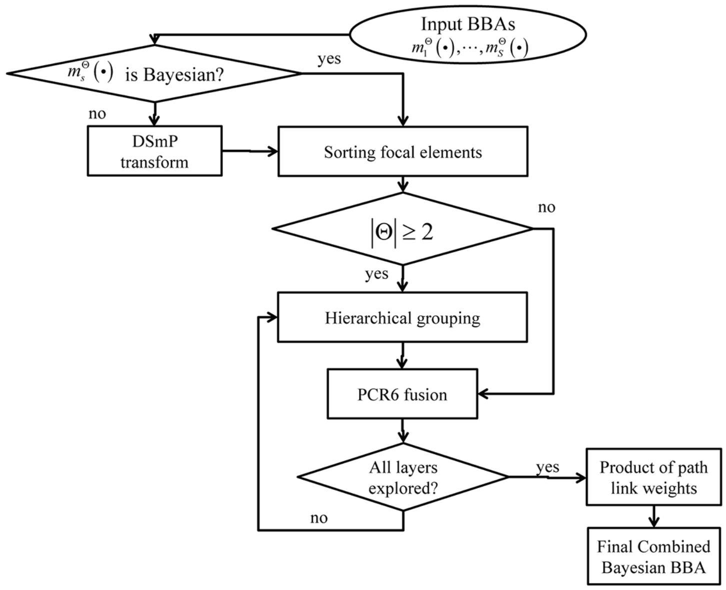


Fig 3. Modified rigid coarsening of FoD for fusion.

https://doi.org/10.1371/journal.pone.0189703.g003

Algorithm 1: Modified Rigid Coarsening Method

Input: All original BBAs $m_1^\Theta(\cdot),\cdots,m_s^\Theta(\cdot)$, $s = 1, 2, \cdots, S$
Output: The final combined BBA $m^\Theta(\cdot)$
if Compound focal elements in Θ: $\theta_i \cup \theta_j \neq \emptyset$ or $\theta_i \cap \theta_j \neq \emptyset$ **then**
  Probabilistic transformation: $DSmP(m_1^\Theta(\cdot)), DSmP(m_2^\Theta(\cdot)),\cdots,DSmP(m_s^\Theta(\cdot))$
end
for $i \leq n$ **do**
  for $s \leq S$ **do**
    Calculate $D_{1-S}(\theta_i) \triangleq \frac{1}{2}\sum_{i=1}^{S}\sum_{j=1}^{S}|d_{BI}(m_i^\Theta(\cdot),m_{\theta_i}(.)) - d_{BI}(m_j^\Theta(\cdot),m_{\theta_i}(.))|$
  end
end
for $i \leq n$ **do**
  Sorting $D_{1-S}(\theta_i)$ in an ascending order.
end
while $|\Theta| \geq 2$ **do**
  if n is an even number **then**
    $m^{\Omega_l}(\omega_1) = \sum_{j=1,\ldots,\frac{n}{2}} m^\Theta(\theta_j)$;
    $m^{\Omega_l}(\omega_2) = \sum_{j=\frac{n}{2}+1,\ldots,n} m^\Theta(\theta_j)$;
  else
    $m^{\Omega_l}(\omega_1) = \sum_{j=1,\ldots,[\frac{n}{2}+1]} m^\Theta(\theta_j)$;
    $m^{\Omega_l}(\omega_2) = \sum_{j=[\frac{n}{2}+1]+1,\ldots,n} m^\Theta(\theta_j)$;
  end
  Then connection weights λ is calculated: $PCR6(m^\Omega(\omega_1), m^\Omega(\omega_2))$
end
foreach focal element θ_i, $i \in 1, \cdots, n$ **do**
  $m^\Theta(\theta_i)$ equals to the product of path link weights from root to terminal nodes.
end

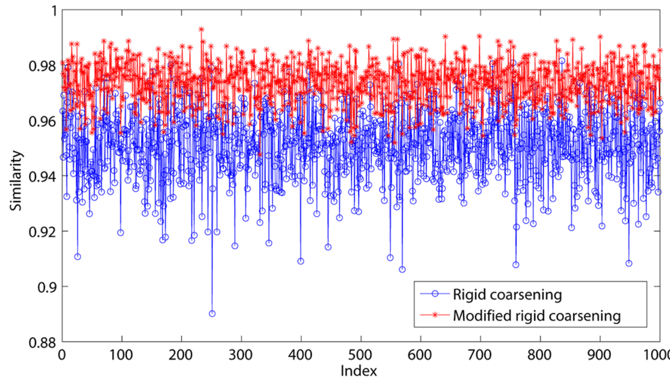


Fig 4. Accuracy comparisons between MRC and PCR6 (Only Singletons).

https://doi.org/10.1371/journal.pone.0189703.g004

Simulation considering accuracy and computational efficiency

- Accuracy:
  Assuming that the FoD is $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{14}, \theta_{15}, \theta_{16}, \theta_{17}, \theta_{18}, \theta_{19}, \theta_{20}\}$, then 1000 BBAs are randomly generated to be fused with three methods: modified rigid coarsening, rigid coarsening and also PCR6. And then distances of fusion results are computed using d_{BI} between two pairs: modified rigid coarsening and PCR6; rigid coarsening and PCR6. Comparisons are made in Fig 4, which show the superiority of our new approach proposed in this paper (The average value of the approximation of modified rigid coarsening is **97.5%** and original rigid coarsening is **94.5%**). Here, similarity represents the approximate degree between fusion results using hierarchical approximate method (both rigid and modified rigid coarsening) and PCR6.

- Computational efficiency:
  As we mentioned before, another advantage of the hierarchical combination method is the computational efficiency. Here, two experiments are conducted (*All experiments are implemented on a PC with I3 CPU, Integrated graphics chipsets and 4G DDR*): 1) the number of singletons is unchanged while the number of BBAs to be fused is increasing; 2) the number of BBAs is unchanged while the number of singletons in FoD is increasing. The results are illustrated in Fig 5 and 6. From experiment 1, all these three methods (classical PCR6, rigid coarsening and also modified rigid coarsening) calculate quickly (**less than 1.2s**) even the number of BBAs increases from 100 to 1000. However, such situation deteriorates when the number of focal elements increases. In Fig 6, when the number of focal elements increases to 500, time consumption of three combinations is: PCR6: 20.6857s; modified rigid coarsening: 7.3320s; rigid coarsening: 5.9748s. This phenomenon also proves that it is reasonable to map original FoD to the coarsening FoD, with the aim of reducing the number of focal elements at the time of fusion. But in any case, computing efficiency of rigid coarsening or modified rigid coarsening is still better than PCR6. On the other hand, modified rigid coarsening makes a significant improvement (accuracy) at the expense of parts of the computational efficiency.

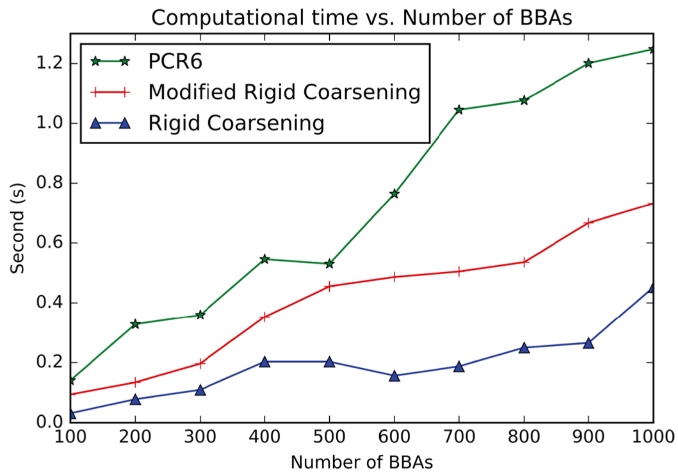


Fig 5. Efficiency comparisons between MRC, RC and PCR6 (With the number of BBAs increasing).

https://doi.org/10.1371/journal.pone.0189703.g005

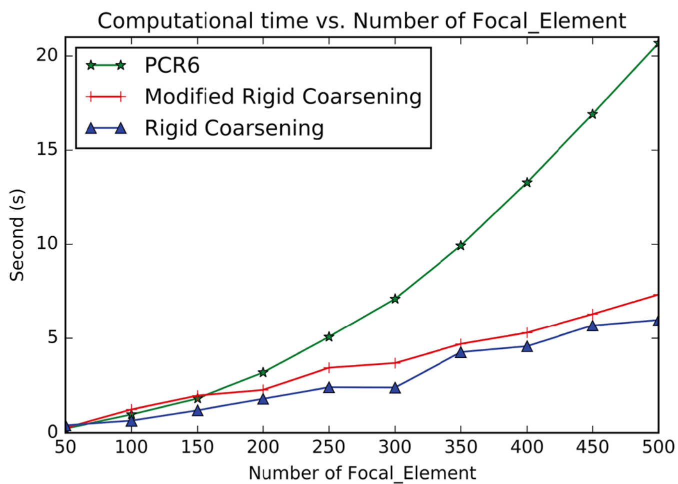


Fig 6. Efficiency comparisons between MRC, RC and PCR6 (With the number of focal elements increasing).

https://doi.org/10.1371/journal.pone.0189703.g006

A recommender system integrating with hierarchical coarsening combination method

In today's e-commerce, online providers often recommend proper goods or services to each consumer based on their personal opinions or preferences [21], [22]. However, it is a tough task to provide appropriate recommendation which may confront several difficulties. One difficulty is that users' preferences are usually characterized as uncertain, imprecise or incomplete [23], [24], which cannot be used directly in RSs. Besides, it is easy to understand that when the more information about user preferences are, the more accurate prediction of RSs will be [25], [26]. But, the problem is that which method we adopt to integrate multi-source uncertain information?

As a general framework for information fusion, DST can not only model uncertain information, but also provide an efficient way to combine multi-source information. These

mentioned features make this theory a wide range of applications [27–29], especially in RSs [23, 25, 30–32]. According to DST, users' comments on products in RSs are described by using mass functions and rules of combination method are used frequently in order to provide appropriate recommendation.

As mentioned in previous sections, both the performances of combination rules in DST or in DSmT suffer from computational complex which is obviously ignored in [23, 25]. Thus, in this paper, modified rigid coarsening method is applicable to combine the imprecise users' preferences in RSs. First, we are required to introduce the relevant knowledge of RSs. Actually, almost all characteristics of RSs have been introduced in [23, 25, 30–32].

First, we give the corresponding representation of the mathematical notation in RSs based on DSmT. RSs usually contain two objects: {*Users*, *Items*}. A set of M users and a set containing N items is respectively denoted by $\mathbf{U} = \{U_1, U_2, \cdots, U_M\}$ and $\mathbf{I} = \{I_1, I_2, \cdots, I_N\}$. Besides, we assume that users can give the corresponding ratings to the items, which include L rating levels ($\Theta = \{\theta_1, \theta_2, \cdots, \theta_L\}$.). Here, L preference levels means multi-level evaluation results. For example, four-levels of user evaluation on the product are {*Excellent*, *Good*, *Fair*, *Poor*}. $r_{i,k}$ means a rating of user U_i on item I_k and a rating matrix $\mathbf{R} = \{r_{i,k}\}$ comprises all the ratings of users on items. It should be noted that $r_{i,k}$ is originally modeled as a mass function $m_{i,k}$: $D^{\Theta} \rightarrow [0, 1]$. Additionally, let $\mathbf{I}_i^R$ and $\mathbf{U}_k^R$ denote the set of items rated by user U_i and the set of users having rated item I_k, respectively.

Contextual information can often be summarized into several genres that significantly affect user's rating of items. Normally, we represent contextual information by a set containing P genres, denoted by $\mathbf{S} = \{S_1, S_2, \cdots, S_P\}$. And each genre S_p, with $1 \leq p \leq P$ contains at most Q groups, denoted by $S_p = \{g_{p,1}, g_{p,2}, \cdots, g_{p,q}, \cdots, g_{p,Q}\}$, $1 \leq q \leq Q$. For a genre $S_p \in \mathbf{S}$, a user $U_i \in \mathbf{U}$ can be interested in several groups and also an item $I_i \in \mathbf{I}$ can belong to one or some groups of this genre, which can be seen in Fig 7.

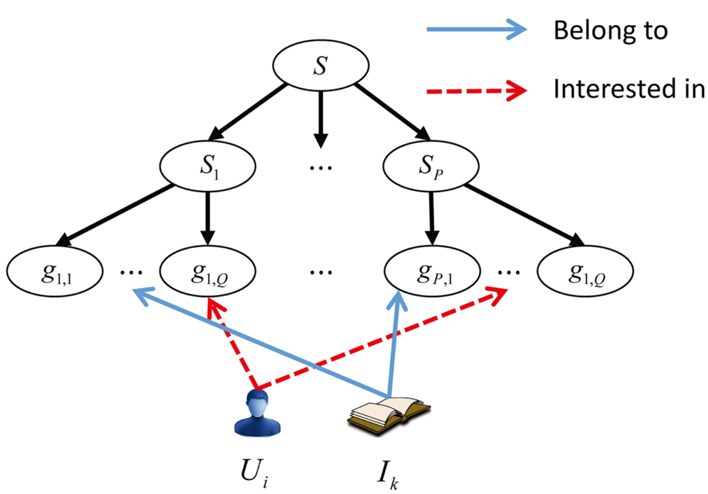


Fig 7. Contextual information.

https://doi.org/10.1371/journal.pone.0189703.g007

Definition 4: *In order to facilitate such expression, two functions* $\kappa(\cdot)$ *and* $\varphi(\cdot)$ *are defined to determine the groups in which user* U_i *is interested and the groups to which item* I_k *belongs, respectively*:

$$\kappa_p : U_i \mapsto \kappa_p(U_i) \subseteq S_p \tag{12}$$

$$\varphi_p : I_k \mapsto \varphi_p(I_k) \subseteq S_p \tag{13}$$

Generally, the main steps of a recommendation system is illustrated in Fig 8, which will be presented in details as follows:

1. DSmT-Modeling Function
   Regarding the DS-partial probability models proposed in [23], the existing ratings $r_{i,k}$, of user U_i on item I_k, are modeled by DSmT-modeling function $M(\cdot)$ in order to transform such hard ratings into the corresponding soft ratings represented as $m_{i,k}$ as below:
   Definition 5:

$$m_{i,k} = \begin{cases} \alpha_{i,k}(1-\sigma_{i,k}), & for \quad A = \theta_l; \\ \frac{1}{2}\alpha_{i,k}\sigma_{i,k}, & for \quad A = B; \\ \frac{1}{2}\alpha_{i,k}\sigma_{i,k}, & for \quad A = C; \\ 1-\alpha_{i,k}, & for \quad A = \Theta; \\ 0, & otherwise. \end{cases} \tag{14}$$

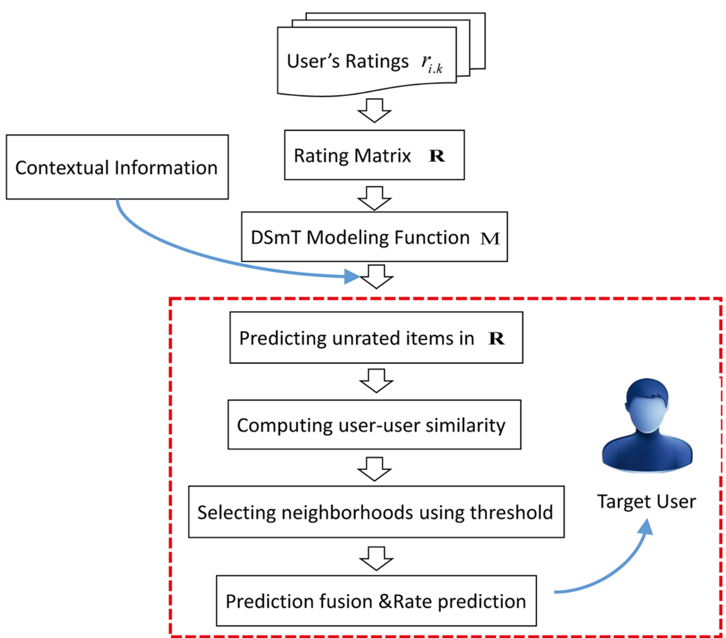


Fig 8. General process of recommendations.

https://doi.org/10.1371/journal.pone.0189703.g008

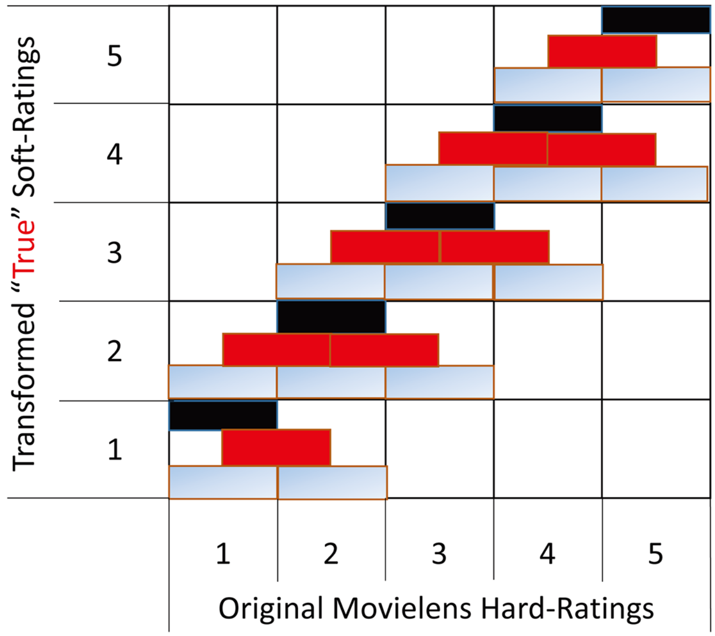


Fig 9. DSmT modeling function.

https://doi.org/10.1371/journal.pone.0189703.g009

with

$$B = \begin{cases} \theta_1 \cup \theta_2, & \text{if} \quad l = 1; \\ \theta_{L-1} \cup \theta_L, & \text{if} \quad l = L; \\ \theta_{l-1} \cup \theta_l \cup \theta_{l+1}, & \text{otherwise}. \end{cases}$$

$$C = \begin{cases} \theta_1 \cap \theta_2, & \text{if} \quad l = 1; \\ \theta_{L-1} \cap \theta_L, & \text{if} \quad l = L; \\ (\theta_{l-1} \cap \theta_l, \theta_l \cap \theta_{l+1}), & \text{otherwise}. \end{cases}$$

where $\alpha_{i,k} \in [0, 1]$ *and* $\sigma_{i,k}$ *are a trust factor and a dispersion factor, respectively* [23]. Referring to the partial probability model analysis in [23], we also give the corresponding user profiles which can be seen in Fig 9. Compared to [23], the difference is that we not only consider the union (*black and gray rectangle*), but also consider the intersection (*red rectangle*) of the hard ratings, which is also the distinction between DS theory and DSmT theory.

Lemma 1: Referring to Definition 5, we can also generate the relative refined BBA in the framework of DS theory:

$$m_{i,k}^{Refined} = \begin{cases} \alpha_{i,k}(1-\sigma_{i,k}), & for \quad A=\theta_l; \\ \alpha_{i,k}\sigma_{i,k}, & for \quad A=B; \\ 1-\alpha_{i,k}, & for \quad A=\Theta; \\ 0, & otherwise. \end{cases} \tag{15}$$

with

$$B = \begin{cases} \theta_1 \cup \theta_2, & if \quad l=1; \\ \theta_{L-1} \cup \theta_L, & if \quad l=L; \\ \theta_{l-1} \cup \theta_l \cup \theta_{l+1}, & otherwise. \end{cases}$$

where $\alpha_{i,k} \in [0, 1]$ and $\sigma_{i,k}$ are a trust factor and a dispersion factor, respectively [23]. After soft ratings are generated, DSmP [33] is applied to decouple non-Bayesian $m_{i,k}$, since the hierarchical fusion algorithm is currently just available for Bayesian BBAs.
Definition 6: *DSmP is a new generalized pignistic transformation defined by* $DSmP_\varepsilon(\emptyset) = 0$ *and for any singleton* $\theta_i \in \Theta$ *by*

$$DSmP_\varepsilon(\theta_i) \triangleq m(\theta_i) + (m(\theta_i) + \varepsilon) \times \left\{ \sum_{A\in 2^\Theta, \theta_N \subset A, |A|\geq 2} \frac{m(A)}{\sum_{B\in 2^\Theta, B\subset A, |B|=1} m(B) + \varepsilon \cdot |A|} \right\} \tag{16}$$

As shown in [33], DSmP makes a remarkable improvement compared with BetP and CuzzP, since a more judicious redistribution of the ignorance masses to the singletons has been adopted by DSmP.

2. Predicting unrated items:
Assuming that users who are keen on the similar groups tend to have common preferences. In this RS, it is necessary to predict the unrated items first. Considering a group $g_{p,q} \in S_p$ with $g_{p,q} \in \varphi(I_k)$, every soft rating, $m_{i,k}$, of user U_i, who is keen on group $g_{p,q}$, on item I_k is regarded as a block of common preference for group $g_{p,q}$. Thus, $G_{m_{p,q,k}}: D^\Theta \to [0, 1]$ which represents all users' group preferences on item I_k regarding group $g_{p,q}$, is computed as follows

$$G_{m_{p,q,k}} = \bigoplus_{\{j | I_k \in \mathbf{I}_j^R, g_{p,q} \in \kappa_p(U_j), g_{p,q} \in \varphi_p(I_k)\}} m_{j,k} \tag{17}$$

Supposing that item I_k has not been rated by user U_i, it usually contains three steps to generate unprovided rating $r_{i,k}$ of user U_i which are shown as below

- Step one: Considering a concept S_p, for each group $g_{p,q} \in \kappa_p(U_i) \cap \varphi_p(I_k)$, it is assumed that all users' group preferences on item I_k regarding group $g_{p,q}$ imply common preference of U_i on I_k regarding group $g_{p,q}$. Furthermore, this group preference is regarded as a piece of user U_i's concept preference on item I_k regarding concept S_p. Therefore, concept preference of user U_i on item I_k regarding concept S_p, denoted by mass function $S_{m_{p,q,k}}: D^\Theta \to$

[0, 1], can be computed as below

$$S_{m_{p,q,k}} = \bigoplus_{\{q|g_{p,q}\in\kappa_p(U_i),g_{p,q}\in\varphi_p(I_k)\}} G_{m_{p,q,k}} \tag{18}$$

- Step two: If there exists at least one common group in concept S_p which item I_k belongs to and also user U_i is interested in, then U_i's concept preference on item I_k regarding concept S_p is regarded as a piece of context preference. Therefore, this user's contextual preference on item I_k, denoted by mass function $S_{m_{i,k}}$: $D^\Theta \to [0, 1]$, is achieved as follows

$$S_{m_{i,k}} = \bigoplus_{p=1,\cdots,P} S_{m_{p,i,k}} \tag{19}$$

- Step three: Context preference of U_i on item I_k is assigned to unprovided rating $\bar{m}_{i,k}$ as below

$$\bar{m}_{i,k} = S_{m_{i,k}} \tag{20}$$

So far, all unprovided ratings are predicted in this RS. Subsequently, user-user similarities are computed depending on both provided and predicted ratings in the following steps.

3. Computing user-user similarities:
Here, we use the distance measure proposed in [34] to calculate distances between two users U_i and U_j with $i \neq j$, which is defined as below

$$D(U_i, U_j) \quad = \sum_{k=1}^{N}\left(\ln \max_{\theta\in\Theta} \frac{m_{j,k}(\theta)}{m_{i,k}(\theta)} - \ln \min_{\theta\in\Theta} \frac{m_{j,k}(\theta)}{m_{i,k}(\theta)}\right) \tag{21}$$

where $m_{i,k}$ and $m_{j,k}$ are the soft ratings of user U_i and user U_j on item I_k respectively. Afterwards, the degree of similarity between U_i and U_j, denoted by $s_{i,j}$, is calculated as follows

$$s_{i,j} = e^{-\gamma\times D(U_i,U_j)}, \mathit{where} \quad \gamma \in (0, \infty). \tag{22}$$

Obviously, if the value of $s_{i,j}$ is high, it means the user U_i and user U_j are very close, and vice versa. Eventually, a mathematical matrix $\mathbf{S} = \{s_{i,j}|U_i, U_j \in \mathbf{U}, i \neq j\}$ is employed to represent the similarities among all users.

4. Selecting neighbors based on user-user similarities:
Taking into account an active user U_i, for each unrated item I_k by user U_i, a set containing K nearest neighborhoods, denoted by $\mathfrak{N}_{i,k}$, is chosen by using the method proposed in [35]. Two simple steps of this method are shown below

- Step one: the process of such selection depends on two criteria: 1. Those users who rated I_k and 2. The corresponding user-user similarities with user U_i are equal or greater than the threshold τ. $\mathfrak{N}_{i,k}$ denotes the selected set, which is acquired as follows:

$$\mathfrak{N}_{i,k} = \{U_j \in \mathbf{U}|I_k \in I_j^R, s_{i,j} \geq \tau\} \tag{23}$$

- Step two: all of members in $\mathfrak{N}_{i,k}$ is descending sorted by $s_{i,j}$ and top K members are selected as the neighborhood set $\mathfrak{N}_{i,k}$.

5. Estimating ratings according to neighborhoods:
Supposing that item I_k has not been rated by user U_i. The predicted rating of U_i on item I_k is denoted as $\hat{m}_{i,k}$. Thus, $\hat{m}_{i,k}$ is calculated according to the ratings of user U_i's nearest users. Mathematically, $\hat{m}_{i,k}$ is given as below

$$\hat{m}_{i,k} = m_{i,k} \bigoplus \tilde{m}_{i,k} \tag{24}$$

where $\tilde{m}_{i,k}$ is the mass regarding the neighborhoods' whole preference in the set Eq (23) on item I_k. Considering user $U_i \in \mathfrak{N}_{i,k}$, and supposing that $s_{i,j}$ is the similarity between user U_i and user U_j. We use a discount rate $1 - s_{i,j}$ to discount the rating of user U_j on item I_k. Therefore, $\tilde{m}_{i,k}$ is:

$$\tilde{m}_{i,k} = \bigoplus_{\{j|U_j \in \mathfrak{N}_{i,k}\}} \dot{m}_{j,k}^{s_{i,j}} \tag{25}$$

$$where \quad \dot{m}_{j,k}^{s_{i,j}} = \begin{cases} s_{i,j} \times m_{j,k}(A), & for\ A \subset \Theta; \\ s_{i,j} \times m_{j,k}(\Theta) + (1 - s_{i,j}), & if\ \ A = \Theta. \end{cases}$$

6. Generating recommendations:
In order to generate appropriate recommendations for the candidate user U_i, predicted ratings of U_i on all unprovided items are sorted, and then based on the sorted list, the appropriate recommendations are generated.

Experiments

To evaluate the performance of modified rigid coarsening in precision of recommendation and computational time, original rigid coarsening method and also classical PCR6 combination method are selected to be regarded as baselines. Besides, we use DS-MAE [23] to measure the precision of recommendations.

Definition 7: *DS-MAE is mathematically given as follows*

$$DS\text{-}MAE(\theta_j) = \frac{1}{|D_j|} \sum_{(i,k) \in D_j, \theta_j \in \Theta} |\hat{m}_{i,k}(\theta_j) - M(\theta_j)| \tag{26}$$

where D_j is the testing set identifying the user-item pairs whose true rating is $\theta_j \in \Theta$.

Those specific users' interested information about genres is unknown. Thus, we define a rule that if a user has rated an item then this user is interested in all genres to which the item belongs.

1. Experiment One:
Movielens (http://grouplens.org/datasets/movielens) is a movie recommendation dataset widely used for benchmarking process. There are nearly 100,000 hard ratings on 19 different types of movies (Action, Comedy and so on). The domain of such rating given in Movielens includes 5 levels, denoted as $\Theta = \{1, 2, 3, 4, 5\}$,. At the same time, each user is required to evaluate at least 20 movies, so as to ensure adequate rating information.
The relevant parameters used in RSs are setted: $\gamma = 10^{-4}$ and $\forall(i, k)\{\alpha_{i,k}, \sigma_{i,k}\} = \{0.9, 2/3\}$. However, Setting parameter τ to be a fixed value is obviously unreasonable because the similarity between two users is quite different when using different combination methods. Hence, in this paper, the value of parameter τ will not be setted in advance. Instead, it is

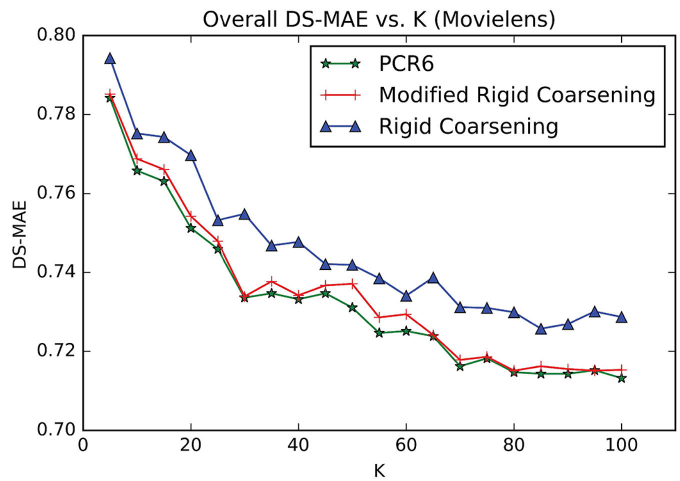


Fig 10. Overall DS-MAE between three combination methods. (Movielens).

https://doi.org/10.1371/journal.pone.0189703.g010

determined based on the similarity in matrix **S**. Specifically, the highest value of top 30% in **S** is selected for τ.

Additionally, we adopt the robust strategy of 10-fold cross validation to conduct experiments, which is widely applied in experimental verification. Specific steps are as follows: original ratings in Movielens are first randomly divided into 10-folds and the experiments are thus carried out 10 times: in each sub-experiment, nine tenths of the ratings are chosen as training data and the remaining ratings are regarded as testing data. It's worth noting that all results illustrated in the following experiments are the average values of 10 times.

Fig 10 demonstrates the values of overall *DS-MAE* varying with changing neighborhood size K. And the smaller values of DS-MAE indicate the better ones. As can be seen in Fig 10, with $K \leq 70$ performances of the three methods increase sharply as well as being the same as each other. With $K \geq 70$, performances of both methods become stable. Especially, performance of modified rigid coarsening method is very close to classical PCR6 rules. However, original rigid coarsening is slightly worse than the other two algorithms.

Fig 11 depicts the computational time varying with changing neighborhood size K. In this figure, the time taken by hierarchical coarsening combination methods (both rigid coarsening and modified rigid coarsening method) is quite faster compared to classical PCR6. Besides, modified rigid coarsening is relatively slower than original rigid coarsening. All these results illustrate that modified rigid coarsening method sacrifices some of the computational efficiency, in exchange for upgrading the accuracy of approximation.

2. Experiment Two:
   Flixster (http://www.cs.ubc.ca/jamalim/datasets/) is a classical recommendation dataset which nearly contains 535013 hard ratings on 19 different types of movies (Drama, Comedy

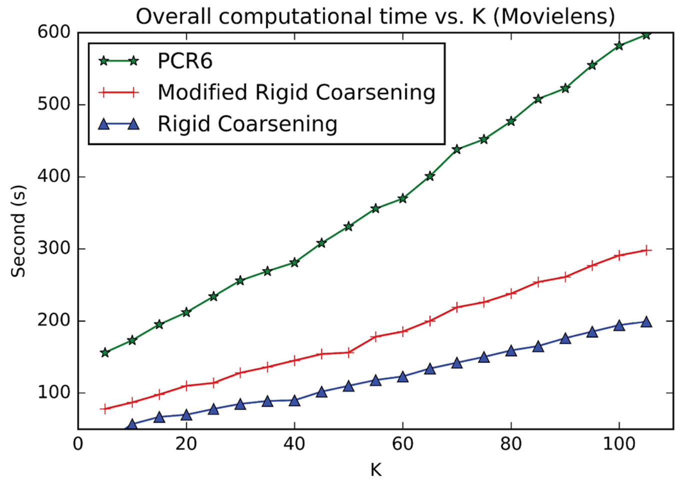


Fig 11. Overall computational time between three combination methods. (Movielens).

https://doi.org/10.1371/journal.pone.0189703.g011

and so on). The domain of such rating given in Flixster includes 10 levels, denoted as $\Theta = \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0\}$,. At the same time, each user is required to evaluate at least 15 movies, so as to ensure adequate rating information. The relevant parameters used in RSs are setted: $\gamma = 10^{-4}$ and $\forall(i, k)\{\alpha_{i,k}, \sigma_{i,k}\} = \{0.9, 2/3\}$. However, Setting parameter τ to be a fixed value is obviously unreasonable because the similarity between two users is quite different when using different combination methods. Hence, in this paper, the value of parameter τ will not be setted in advance. Instead, it is determined based on the similarity in matrix **S**. Specifically, the highest value of top 50% in **S** is selected for τ.

Fig 12 demonstrates the values of overall *DS-MAE* varying with changing neighborhood size *K*. And the smaller values of DS-MAE indicate the better ones. As can be seen in Fig 12, we can get a similar result to the previous data set(Movielens). Especially, performance of modified rigid coarsening method is in the middle of the comparison methods. However, original rigid coarsening is worse than the other two algorithms. Fig 13 depicts the computational time varying with changing neighborhood size *K*. From this figure, we can also get the same conclusion that the time taken by hierarchical coarsening combination methods (both rigid coarsening and modified rigid coarsening method) is quite faster compared to classical PCR6.

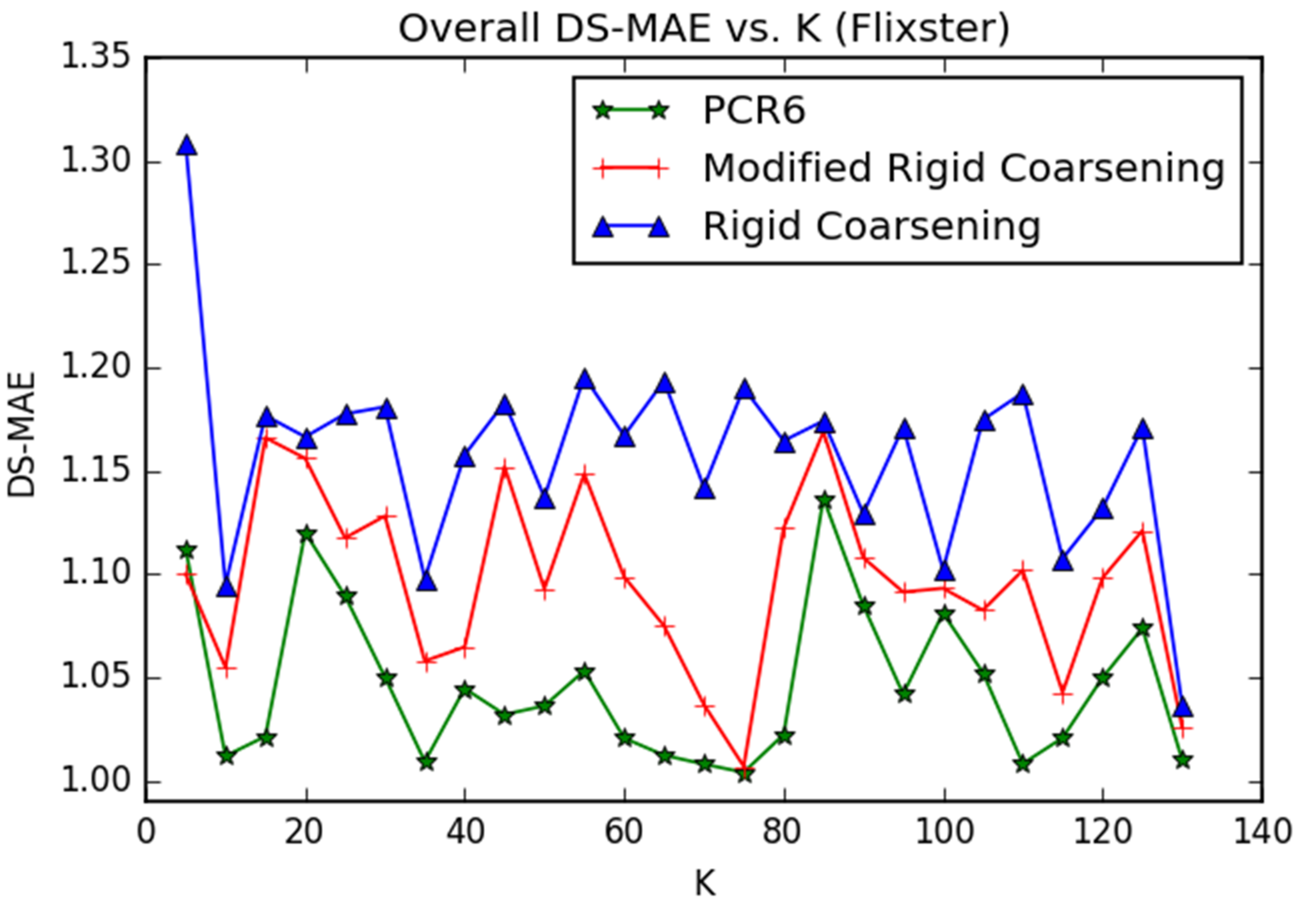


Fig 12. Overall DS-MAE between three combination method. (Flixster).

https://doi.org/10.1371/journal.pone.0189703.g012

Conclusion

In this paper, we propose a new combination method, called modified rigid coarsening method. This new method can map the original refined FoD to the new coarsening FoD in the process of combination. Compared to traditional fusion method PCR6 in DSmT, this approach can not only reduce computational complexity, but also ensure high approximation accuracy. Besides, in order to verify the practicality of our approach, we apply this approach to fuse soft ratings in RSs. To be specific, user preferences are first transformed by DSmT-partial probability model to accurately represent uncertain information. Then, information about user preferences from different sources can be easily combined. In the future work, more helpful information will be mined to discern focal element in FoD so as to improve the accuracy of approximation and more data sets will be applied.

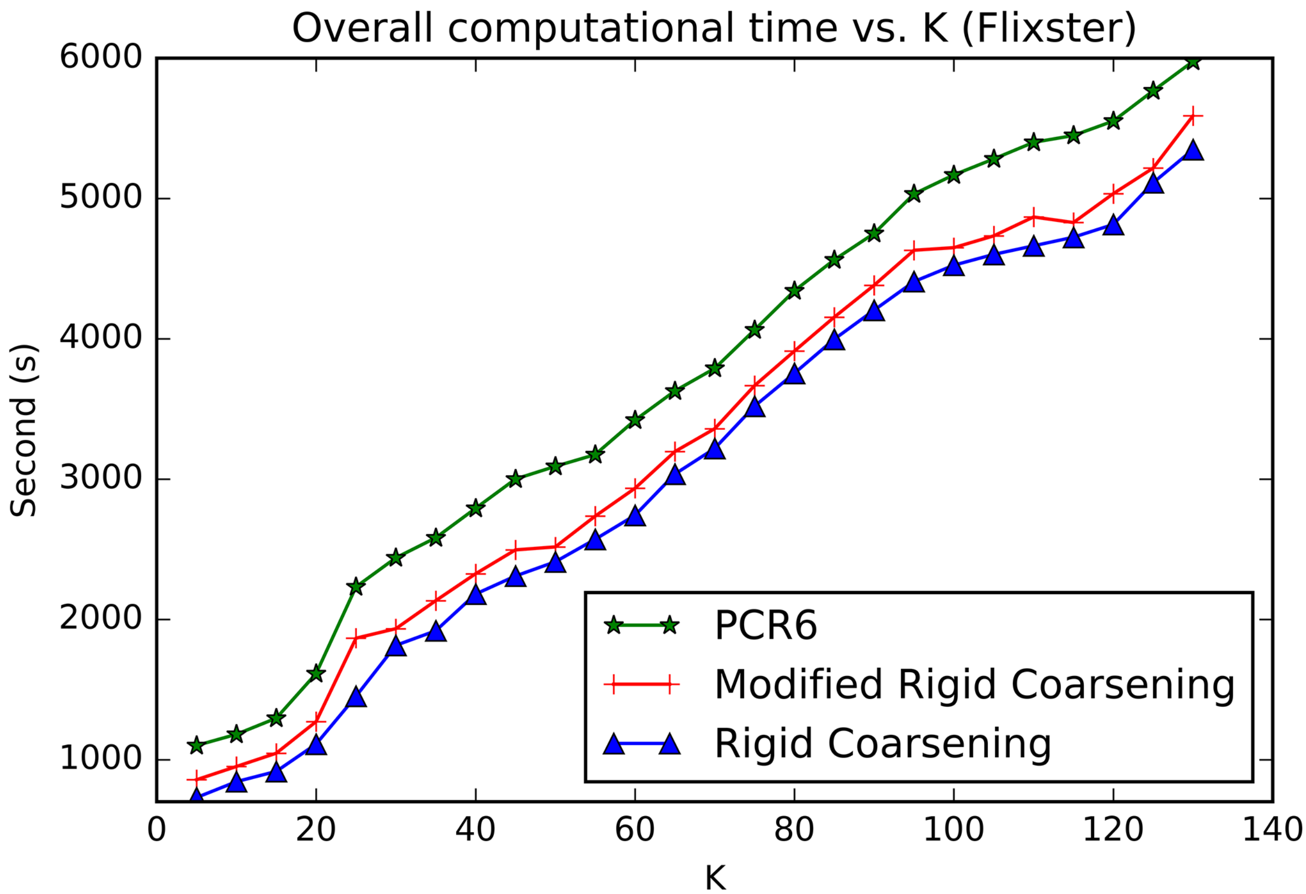


Fig 13. Overall computational time between three combination methods. (Flixster).

https://doi.org/10.1371/journal.pone.0189703.g013

Supporting information

S1 File. The compressed file package of all datasets used in this paper.
(RAR)

S1 Appendix. Proof of D_{1-S} in Eq (9).
(DOCX)

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61573097, 91748106, in part by Key Laboratory of Integrated Automation of Process Industry (PAL-N201704), in part by the Qing Lan Project and Six Major Top-talent Plan, and in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions. The authors thank the reviewers and editors for giving valuable comments, which are very helpful for improving this manuscript.

Author Contributions

Investigation: Yilin Dong.

Methodology: Yilin Dong, Xinde Li.

Supervision: Xinde Li, Yihai Liu.

Validation: Yilin Dong.

Visualization: Yilin Dong.

Writing – original draft: Yilin Dong.

References

1. Shafer G. A mathematical theory of evidence. Princeton Univ. Press, 1976.
2. Dempster A. Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics. 1967; 38:325–339. https://doi.org/10.1214/aoms/1177698950
3. Jiang W, Wang S, Liu X, Zheng H, Wei B. Evidence conflict measure based on OWA operator in open world. PLoS ONE. 2017; 12(5):e0177828. http://doi.org/10.1371/journal.pone.0177828 PMID: 28542271
4. Smets P. Practical uses of belief functions. In K.B. Laskey and H. Prade Editors, 15th Conf. on Uncertainty in Artificial Intelligence, pp. 612–621, Stockholm, Sweden, 1999.
5. Dezert J. Foundations for a new theory of plausible and paradoxical reasoning. Information & Security: An Int. Journal. 2002:9.
6. Smarandache F, Dezert J (Editors). Advances and applications of DSmT for information fusion. American Research Press, Rehoboth, NM, U.S.A., Vol. 1–4, 2004–2015. Available at webpage 2 of [7].
7. http://www.onera.fr/staff/jean-dezert
8. Kennes R. Computational aspects of the Möbius transform of graphs, IEEE Trans. on SMC. 1992; 22:201–223.
9. Shafer G, Logan R. Implementing Dempster's rule for hierarchical evidence, Artificial Intelligence. 1987; 33:271–298. https://doi.org/10.1016/0004-3702(87)90040-3
10. Yang Y, Liu YL. Iterative approximation of basic belief assignment based on distance of evidence. PLoS ONE. 2016; 11(2):e0147799. https://doi.org/10.1371/journal.pone.0147799 PMID: 26829403
11. Denœux T. Inter and outer approximation of belief structures using a hierarchical clustering approach. Int. J. of Uncertainty, Fuzziness and Knowledge-Based System. 2001; 9(4):437–460. https://doi.org/10.1142/S0218488501000880
12. Yang Y, Han DQ, Han CZ, Cao F. A novel approximation of basic probability assignment based on rank-level fusion. Chinese Journal of Aeronautics. 2013; 26(4):993–999. https://doi.org/10.1016/j.cja.2013.04.061
13. Han DQ, Yang Y, Dezert J. Two novel methods of BBA approximation based on focal element redundancy. Proc. of Fusion 2015, Washington, D.C., USA, July 2015.
14. Li MZ, Zhang Q, Deng Y. A New Probability Transformation Based on the Ordered Visibility Graph. International Journal of Intelligent Systems. 2016; 31(1):44–67. https://doi.org/10.1002/int.21744
15. Dong YL, Li XD, Dezert J. A Hierarchical Flexible Coarsening Method to Combine BBAs in Probabilities, accepted in 20th International Conference on Information Fusion (Fusion 2017), Xi'an, China, July 10-13, 2017.
16. Smets P. Analyzing the combination of conflicting belief functions, in Information Fusion.2006; 8:387–412. https://doi.org/10.1016/j.inffus.2006.04.003
17. Li XD, Dezert J, Huang XH, Meng ZD, Wu XJ. A fast approximate reasoning method in hierarchical DSmT (A). Acta Electronica Sinica.2010; 38(11):2567–2572.
18. Li XD, Yang WD, Wu XJ, Dezert J. A fast approximate reasoning method in hierarchical DSmT (B). Acta Electronica Sinica.2011; 39(3A):32–36.
19. Li XD, Yang WD, Dezert J. A fast approximate reasoning method in hierarchical DSmT (C). J. Huazhong Univ. of Sci. and Tech. (Natural Science Edition).2011; 39:151–156.
20. Han DQ, Dezert J and Yang Y. Belief Interval-Based Distance Measures in the Theory of Belief Functions. IEEE Transactions on Systems, Man and Cybernetics: Systems.2016:1–18.

21. Bobadilla J, Orrega F, Hernando A and Gutierrez A. Recommender systems survey. Knowledge-Based System. 2013; 46:109–132. https://doi.org/10.1016/j.knosys.2013.03.012
22. Chen T. Ubiquitous Hotel Recommendation Using a Fuzzy-Weighted-Average and Backpropagation-Network Approach. International Journal of Intelligent Systems. 2017; 32(4):00–00. https://doi.org/10.1002/int.21863
23. Wichramarathne TL, Premaratne K, Kubat M, Jayaweera DT. CoFiDS: a belief-theoretic approach for automated collaborative filtering. IEEE Transaction on Knowledge and Data Engineering. 2011; 23 (2):175–189. https://doi.org/10.1109/TKDE.2010.88
24. Ladyzynski P and Grzegorzewski P. Vague preferences in recommender systems. Expert Systems With Applications. 2015; 42(24):9402–9411. https://doi.org/10.1016/j.eswa.2015.08.006
25. Nguyen VD, Huynh VN. Two-probabilities focused combination in recommender systems. International Journal of Approximate Reasoning.2017; 80:225–238. https://doi.org/10.1016/j.ijar.2016.09.005
26. Bagher RC, Hassanpour H and Mashayekhi H. User trends modeling for a content-based recommender system. Expert Systems With Applications. 2017; 87:209–219. https://doi.org/10.1016/j.eswa.2017.06.020
27. Denoeux T. Maximum likelihood estimation from uncertain data in the belief function framework. IEEE Transaction on Knowledge and Data Engineering.2013; 25(1):119–130. https://doi.org/10.1109/TKDE.2011.201
28. Kanjanararakul O, Sriboonchitta S, Denoeux T. Forecasting using belief functions: an application to marketing econometrics. International Journal of Approximate Reasoning. 2014; 55(5): 1113–1128. https://doi.org/10.1016/j.ijar.2014.01.005
29. Masson M, Denoeux T. Ensemble clustering in the belief functions framework. International Journal of Approximate Reasoning. 2011; 52(1): 92–109. https://doi.org/10.1016/j.ijar.2010.04.007
30. Troiano L, Rodriguez-Muniz LJ, Diaz J. Discovering user preferences using Dempster-Shafer Theory. Fuzzy Sets Systems.2015; 278: 98–117. https://doi.org/10.1016/j.fss.2015.06.004
31. Nguyen VD, Huynh VN. A reliably weighted collaborative filtering system. ECSQARU 2015:429–439.
32. Jglesias J, Bernardos AM, Casar JR. An evidential and context-aware recommendation strategy to enhance interactions with smart spaces. HAIS 2013:242–251.
33. Dezert J, Smarandache F. A new probabilistic transformation of belief mass assignment. In Proc. of 11th Int. Conf. on Information Fusion, Cologne, Germany, pp. 1-8, June-July 2008.
34. Chan H, Darwiche A. A distance measure for bounding probabilistic belief change. International Journal of Approximate Reasoning.2005;38(2): 149–174.
35. Herlocker JI, Konstan JR, Borchers A and Riedl J. An algorithmic framework for performing collaborative filtering. SIGIR'99, ACM, 1999: 230–237.